JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1998, **31,** 357–374 NUMBER 3 (FALL 1998)

EFFECTS OF DIFFERENTIAL REINFORCEMENT ON THE GENERALIZATION OF A REPLACEMENT MAND IN THREE CHILDREN WITH SEVERE LANGUAGE DELAYS

ERIK DRASGOW

UNIVERSITY OF SOUTH CAROLINA

AND

JAMES W. HALLE AND MICHAELENE M. OSTROSKY

UNIVERSITY OF ILLINOIS AT URBANA–CHAMPAIGN

We investigated variables that may influence the generalization of a replacement mand in 3 young children with severe language delays. A multiple baseline design consisting of one stimulus class of manding opportunities that we arbitrarily divided into three categories (i.e., food, toys, and events) was used for each child. During baseline probes, all children manded mainly by reaching, grabbing, or leading. We then taught each child a replacement mand using a single member of the stimulus class. Acquisition of the replacement mand occurred under highly restricted conditions in a setting that was completely isolated from the generalization settings. Postacquisition probes revealed almost exclusive use of old manding forms. Subsequently, extinction of the old forms and reinforcement of the replacement mand were introduced in a sequential fashion. Two children manifested a substantial increase, and 1 child displayed a moderate increase in the occurrence of the replacement mand (i.e., generalization occurred). These results suggest that a differential reinforcement procedure can alter the probability of the occurrence of response class members across a variety of stimulus conditions.

DESCRIPTORS: response class, language, differential reinforcement, generalization

Generalization of new language targets has been a frequently sought-after but highly elusive goal for students with severe disabilities (e.g., Bryen & Joyce, 1985; Halle & Holt, 1991; Harris, 1975; Kaczmarek, 1990; Stokes & Baer, 1977). Over time, methods have changed from therapy-based communication intervention characterized by one-to-one teaching in contexts isolated from natural settings (e.g., Bricker & Bricker, 1970; Guess, Sailor, & Baer, 1974; Schiefelbusch, 1978; Schiefelbusch & Lloyd, 1974) to more naturalistic approaches that capitalize on normally occurring opportunities to teach language in a variety of everyday contexts (e.g., Chadsey-Rusch, Drasgow, Reinoehl, Halle, & Collet-Klingenberg, 1993; Halle, 1982; Halle, Marshall, & Spradlin, 1979, Hart & Rogers-Warren, 1978; Kaiser, Yoder, & Keetz, 1992; Koegel, O'Dell, & Koegel, 1987; Warren & Rogers-Warren, 1985). Despite improvements in language teaching outcomes, generalization to other settings and situations still may be limited

This study was completed as part of the requirements for the first author's PhD degree in Special Education while he was a student at the University of Illinois. This study was supported in whole or in part by the U.S. Department of Education Office of Special Education Programs, Grant H024H20011. The opinions expressed herein, however, do not necessarily reflect the position or policy of the U.S. Department of Education.

The authors thank Ann, Kate, Tom, their families, and Susan Messer for their participation in this study. We also thank Edward G. Carr and Joseph E. Spradlin for their helpful comments on an earlier draft of this manuscript, and Ed Holland for his technical assistance.

Correspondence and requests for reprints should be addressed to Erik Drasgow, University of South Carolina, 235-I Wardlaw, Columbia, South Carolina 29208.

or inconsistent (e.g., Hemmeter, Ault, Collins, & Meyer, 1996; Horner, Bellamy, & Colvin, 1984; Losardo & Bricker, 1994; Reichle & Sigafoos, 1991).

One potential reason for the limited generalization of new responses (e.g., the word or sign for "want") is that students may have an extant communication topography (e.g., leading, grabbing) that serves an identical function and thus interferes with the generalization of the new response (Drasgow & Halle, 1995). Several investigators have examined this possibility within the context of manding (e.g., Billingsley & Neel, 1985; Carr & Kemp, 1989; Drasgow, Halle, Ostrosky, & Harbers, 1996). For example, Carr and Kemp (1989) determined that 4 young nonverbal children used leading to mand. These researchers then taught pointing as a replacement mand and subsequently increased its probability of generalized occurrence relative to leading. With replacement topographies, a stimulus class already controls a response (or set of responses), and the goal of intervention is to add an alternative response that will become the most probable among response options. Generalization of replacement topographies may differ from generalization of new topographies because replacement topographies interact in complex ways with other existing options in a response class.

A response class consists of behaviors that are topographically different but produce the same effect on the environment (Carr, 1988; Catania, 1998; Johnston & Pennypacker, 1993). There are several dimensions of reinforcement that may affect the probability of the emission of any particular member in a response class. One of these is related to the consequences, or the schedule of reinforcement associated with each individual member (Davison & McCarthy, 1988; Herrnstein, 1970). That is, an increase in the reinforcement of one response is likely to decrease the rate of other responses (Catania, 1966; Parrish, Cataldo, Kolko, Neef, & Egel, 1986). Another variable that may affect responding is the efficiency of each response class member. Response efficiency is determined by criteria that include (a) the amount of physical effort a response requires, (b) the time between a response and reinforcement, and (c) the quality of reinforcement that is received (Friman & Poling, 1995; Horner & Day, 1991; Mace & Roberts, 1993). The schedule of reinforcement and response efficiency may influence the probability of particular responses, and, thus, response classes may assume a hierarchical ordering of individual topographies (Baer, 1982; Lalli, Mace, Wohn, & Livezey, 1995).

Several studies have demonstrated that response class hierarchies can be altered (e.g., Carr & Durand, 1985; Carr & Kemp, 1989; Drasgow et al., 1996; Lalli et al., 1995; Sprague & Horner, 1992). For example, Lalli et al. (1995) determined that a 15-year-old girl with moderate mental disabilities had a response class consisting of screams, aggression, and self-injury that (a) were escape maintained and (b) occurred in a stable sequence (i.e., may have been hierarchically ordered). These researchers then taught a new functionally equivalent vocal response, expanding the existing response class. When the new response was negatively reinforced, the occurrence of other members in the response class was modified.

In the present study, we examined additional mechanisms that affected which member of a manding response class would be emitted under specific stimulus conditions. First, we taught a new manding response to 3 young children with severe disabilities. The new manding response was taught under extremely restricted stimulus conditions that typically reduce the likelihood of generalization. Next, we differentially reinforced the new response and placed competing members of the class on extinction to deter-

mine whether (a) the new response would occur across a variety of stimulus conditions and (b) the probability of occurrence of the new mand would increase over time relative to competing mands. We hypothesized that, following training, the novel mand would be included in the existing manding response class but only as a low-probability member. Placing the existing mands on extinction while reinforcing the new mand would be needed to increase the probability of occurrence of the new mand under generalization conditions.

METHOD

Participants

Three young children who had received a diagnosis from a local childhood disability clinic of either autism or pervasive developmental disorder with severe language delays participated in this study. All lived at home with their families and attended a half-day preschool program.

Ann was 3 years 3 months old when the study began; her expressive language age equivalency was 8 months, and her receptive age equivalency was 4 months (Sequenced Inventory of Communication Development–Revised, SICD–R; Hedrick, Prather, & Tobin, 1984). Ann's parents reported that she was often unresponsive when they talked to her, and her teacher reported that Ann rarely understood language used in class. In addition, her parents and private speech therapist reported that she produced some vocalizations, but the function of these vocalizations was unclear to them. Her parents reported that after several months of private speech therapy, Ann spontaneously had produced the sign for "more" once. Her parents and teachers reported that although she usually played alone, she had several preferred objects and events, such as climbing up the ladder on a swing set to go down the slide and being pushed while on a swing.

Kate was 3 years 4 months old when the study began; no expressive or receptive language age equivalency scores using the SICD–R (Hedrick et al., 1984) could be established for Kate because she lacked a basal score in each area. Kate's parents reported that she was often unresponsive when they talked to her and that she did not show an interest in people around her. In addition, her parents and private speech therapist reported that she produced some vocalizations that they described as "jargon" rather than true words. Her parents and teachers reported that although she usually played alone, she had several preferred objects and events, such as watching herself in a large mirror, running, jumping on a trampoline at home, being pushed while on a swing, and watching "Barney" on the television.

Tom was 3 years old when the study began. In addition to the label of autism, he had a diagnosis of Fragile X syndrome. His expressive and receptive language age equivalencies were both 4 months (SICD–R, Hedrick et al., 1984). Tom's parents reported that he did not talk, but often communicated his needs by approaching them, yelling, flapping his hands, or handing them an object (e.g., his cup). His parents also reported that he enjoyed outdoor events, such as swinging and going for walks, and particularly enjoyed watching different videotapes on television.

Settings

A single training setting and multiple generalization settings were established for each participant. The training setting was independent of, and isolated from, the generalization settings. Ann's training setting was the parent conference room in her preschool. This room was the same size as her classroom and contained some toys kept on shelves, a large table and chairs, and two office desks. Ann's generalization settings were her early childhood classroom, the play-

ground, and the school gym. Ann's preschool program served 8 to 10 children from 3 to 6 years old who had developmental delays or were identified as being at risk for future academic failure. The classroom environment was arranged with various activity centers (e.g., library, play kitchen, sensory table) and had a consistent routine (e.g., short structured times separated by free times and snack time). The generalization settings also included (a) an outdoor playground with swings, a jungle gym, and several tricycles and wagons and (b) an indoor school gym equipped with an indoor climber and slide, tricycles, wagons, and balls. The indoor gym was used on days when the weather was too inclement to go outside.

Kate's training setting was her preschool classroom. Her preschool program served 8 to 10 typically developing children ranging in age from 3 to 5 years old. Training occurred in a corner of the room when other children were either out of the classroom or busy in some other area of the classroom. Kate's generalization settings were in her home. Areas in the home used for the assessment of generalization included a large two-car garage, a pantry with shelves and a washer and a dryer, a kitchen with several counters, an eating area with a table and chairs, and two large family rooms. The family rooms contained toys, a large doll house, an indoor climbing gym, a large television, and typical household furniture.

Tom's training setting was his bedroom on the second floor of his home. The bedroom was approximately 3 m by 3 m, and contained his bed, shelves holding toys and books, a dresser, and a small chair. The generalization settings included the first floor and basement of his home. The basement was a large open area with a swing, a television, a padded mat, and various toys. The first floor consisted of a kitchen with a cooking area and a separate area with a table and chairs, a living room with furniture, a playroom with a television and toys, and a bathroom. Tom's mother also provided child care for other children, and the number of these children in the home varied from none to five, depending on the time and day.

Stimulus and Response Classes

We began the study with interviews (Halle, Chadsey-Rusch, Collet-Klingenberg, & Reinoehl, 1992) and direct observations (Bijou, Peterson, & Ault, 1968). First, we interviewed Ann's parents and special education teacher, Kate's mother, and Tom's parents. The purposes of the interviews were (a) to collect information regarding the potential manding response class topographies of each child and (b) to identify highly preferred items and events for each child. Next, we conducted observations using antecedent-behavior-consequence (ABC) recording in Ann's classroom and in Kate's and Tom's homes to verify the presence of communicative topographies that parents and teachers had identified as potential members of a manding response class. These observations took place during times when the items and events that parents had identified as being highly preferred by their child were available. We observed each participant for approximately 1 to 2 hr per day for at least 2 weeks. Based on the frequent and consistent use of particular topographies that occurred during this time and the repeated and consistent ways in which various social partners responded to these topographies, we hypothesized that each participant had a manding response class consisting of at least three distinct topographies (operational definitions are listed in Table 1). Moreover, these topographies consistently occurred in the presence of three relevant stimulus properties that we used to define the stimulus class for each participant: A preferred item or event was present but unavailable, and a listener mediated access to the item or mediated participation in the event.

Table 1
Operational Definitions of Existing Forms

Existing forms	Operational definitions
Leading	Taking an adult by the hand or wrist, guiding the adult's hand or wrist to an object, and then placing the adult's hand on the object.
Reaching	In the presence of an adult, moving one or both hands toward an object, but not making contact with the object. If the participant does not need an adult to mediate access to the object, then this is not a communicative opportunity (e.g., a participant reaches for the coat in his closet, and he needs neither assistance nor permission to get it, then this is not a communicative opportunity because he can get the coat himself).
Grabbing	Reaching for an object and grasping the object while the object is in the grasp of another's hand (e.g., an adult is holding a can of pop in her hand, the child approaches the adult, grasps the can, and either waits or pulls on the can).
Yelling (Tom only)	Emitting a loud, undifferentiated vocalization of 1 or more seconds in the presence of an adult in which the context makes its requesting function clear (e.g., Tom approaches a closed door, on the other side of which is a preferred person or activity, and yells). Yelling during episodes of crying, or yelling concurrently with flopping or diving on the floor, were not included. Yelling could be coded as a separate form or in combination with another form.

Definition and Selection of the New Form

We selected the sign for "please" from American sign language as the new manding form to be taught to all 3 children. It was operationally defined as patting the center of the chest with an open hand two or more times. Acceptable forms included patting with (a) the left hand only, (b) the right hand only, or (c) both hands simultaneously. We opted for a gestural modality because, prior to the study, parents, teachers, and speech therapists had already selected this modality for the 3 children. All parents and teachers reported that they had never seen any of the participants use "please," and it was not observed prior to the initiation of massed trial training.

Recording System and Interobserver Agreement

Two observers simultaneously but independently coded (a) the first communicative form that occurred during baseline probes and (b) all communicative forms that occurred during intervention probes. Agreement was calculated separately for "please" and for the occurrence and sequence of all other communicative forms listed in Table 1. An agreement for "please" was defined as a probe occasion in which both observers recorded its occurrence. In addition, a point-by-point agreement ratio (Kazdin, 1982) was used to determine agreement on the occurrence and sequence of all other communicative forms. Agreement was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100%.

Agreement checks were conducted on 33% (27 of 82) of Ann's baseline probes, with agreement of 96% (26 of 27); on 36% (24 of 66) of Kate's baseline probes, with agreement of 100%; and on 38% (30 of 78) of Tom's baseline probes, with agreement of 100%. Agreement checks were conducted on 41% (19 of 46) of Ann's intervention probes, with agreement of 100% (16 of 16) on the occurrence of "please" and 92% (22 of 24) on the occurrence and sequence of other forms. Agreement checks were conducted on 23% (5 of 22) of Kate's intervention probes, with agreement of 100% (1 of 1) on the occurrence of "please" and 90% (9 of 10) on the occurrence and sequence of

other forms. Agreement checks were conducted on 26% (10 of 39) of Tom's intervention probes, with agreement of 89% (8 of 9) on the occurrence of "please." Agreement on the occurrence and sequence of Tom's other forms was 75% (12 of 16, range 0% to 100%).

Experimental Design and General Procedure

A multiple baseline design (Baer, Wolf, & Risley, 1968) across the stimulus class items was used to study the effects of intervention. For each participant, the multiple baseline consisted of manding opportunities that we arbitrarily divided into the categories of food, toys, and events.

Baseline 1. Baseline 1 consisted of conducting probes during manding opportunities that were embedded across the normal daily routine in each participant's natural environment (i.e., the generalization settings). We defined a probe as the assessment of a participant's manding behavior when one of the foods, toys, or events was available, *and* a social partner was available and necessary to mediate access to the item or event.

The Baseline 1 protocol consisted of first setting up the situation (e.g., at snack time, a probe food item may have been placed in sight but out of reach). Next, the experimenter made himself or herself available for approximately 5 min to fulfill a request by sitting next to the participant, but no extra prompting (e.g., "Do you want something?") occurred. Finally, the experimenter provided the requested item immediately after the first communicative request (i.e., any of the forms listed in Table 1, or the sign for "please"). No verbal praise occurred when the experimenter provided the requested item. After recording the outcome, the probe was then complete. If the participant did not emit any communicative response during the 5 min when the experimenter was available, the probe was discontinued and was re-presented during the next natural opportunity.

Three other conditions were included in the Baseline 1 protocol. First, at least a 10-min latency was required between probes. Second, no more than two consecutive probes were conducted within a single category (e.g., food) on any day. Probes from each category of the multiple baseline were interspersed with probes from the other two categories of the baseline. The first two conditions were easily incorporated into the protocol because the opportunities to conduct probes from different categories were distributed naturally across each participant's normal daily routine. Third, at least three different people (e.g., experimenter, teacher, teacher's aide) were involved in each probe set to prevent the possibility of one person becoming a consistent stimulus during different manding opportunities. We defined a probe set as completing individual probes on each food, toy, and event contained in the stimulus class for each participant.

If any participant began to engage in problem behavior (e.g., screaming or crying) during a probe in any phase of the study, the probe was discontinued. The participant then was taken from the situation and redirected to another activity for 1 min. The participant was returned to the same situation after (a) 1 min *and* (b) the problem behavior had subsided. Upon return, the participant could directly obtain the probe item.

Massed trial training. The purpose of massed trial training was to teach each participant the sign for "please" under restricted stimulus conditions. Training was conducted (a) in a setting independent of the generalization settings, (b) by an individual never associated with the generalization settings, and (c) with a single stimulus that was the same in every training session and that was never available in the generalization settings. The training stimulus was selected from the

stimulus class for each participant that we had identified earlier. The training stimulus was music sticks for Ann, candy for Kate, and crackers for Tom. We selected the training stimulus according to the following criteria: (a) It was part of an existing stimulus class that controlled manding responses; (b) parents, teachers, or both had identified it as highly preferred for that participant; (c) observations confirmed it as highly preferred; and (d) parents, teachers, or both agreed that the training stimulus would be available only in the massed trial training setting. The last criterion was easily met because parents, teachers, or both had already restricted each participant's access to these items, and all agreed that massed trial training allowed limited but acceptable access to these items.

The training protocol consisted of the following steps. First, the participant was seated in a chair with the trainer seated in front of the participant. An individual trial began when the trainer presented the training stimulus by holding it in front of him or placing it near him, but out of the participant's reach. The trainer then waited for the participant to attempt to obtain (e.g., by reaching or by leading) the training stimulus. In the presence of these attempts, the trainer used graduated guidance consisting entirely of physical prompts (e.g., full physical prompts, partial physical prompts) to shape "please." The trainer never gave verbal prompts or praise. The trainer provided the stimulus immediately after a prompted response, and later immediately after an approximation or an independent response. The trainer then waited 4 to 5 s and began the next trial.

When each participant began to make independent approximations of "please," the trainer lightly blocked or held the participant's other hand to prevent any response chaining of existing forms (e.g., leading, reaching) occurring with "please." A training session ended when 10 min had elapsed from the start of the first trial. When the session was over, the trainer returned the participant to the original setting and departed. A massed trial training session consisted of the total number of individual trials that occurred within that session; the number of individual trials could vary across sessions because of our protocol for ending a session (range, 1 to 30 trials).

Massed trial acquisition criterion was set at the unprompted emission of "please" occurring with no other manding forms during the first five trials of a training session over three consecutive training sessions. Although the criterion included only the first five trials of a session, the session was not terminated until 10 min had elapsed from the first trial.

Once criterion was met, massed trial training occurred only 1 day per week. No probes were conducted after training on that day because the temporal proximity of massed trials to a probe might influence the probability of the target response. If a participant failed to maintain 100% unprompted emission of "please" during the weekly massed trial session, we then reinstated massed trial training every day, with no probes in the generalization settings occurring until the participant met the acquisition criterion again.

Baseline probes during massed trial training. One complete probe set in the generalization settings was conducted under Baseline 1 conditions for each participant prior to reaching criterion during massed trial training. The purpose of this probe set was to determine whether responding with "please" in the generalization settings was in any way related to the level of acquisition in the massed trial setting. This probe set was spread out over the acquisition period, with probes occurring on the same day as training. In general, no more than one or two probes were conducted on any day, and they usually were conducted prior to the day's training session. The generalization probes

were distributed across the acquisition period so that some probes (a) occurred during early acquisition (no spontaneous responding with "please" in the massed trial setting), (b) occurred during the middle of acquisition (some spontaneous responding with "please" in the massed trial setting), and (c) occurred during late acquisition (mostly spontaneous responding with "please" in the massed trial setting).

Postacquisition baseline (Baseline 2). After each participant had mastered "please" in the massed trial training setting (i.e., met acquisition criterion), we conducted more probes using the Baseline 1 probe procedure. The purpose of these probes was to assess whether "please" would be the first topography used in the generalization settings. Massed trial training had established "please" as the first and only response in the isolated training setting with a single stimulus, and Baseline 2 probes assessed whether it was the first response in the generalization settings with multiple stimuli. In sum, this procedure consisted of setting up a potential manding opportunity, reinforcing the first mand, and then ending the probe.

Differential reinforcement of alternative behavior (DRA). The DRA condition was introduced in a sequential fashion across the three categories (i.e., food, toys, and events) of the multiple baseline design. The implementation of the DRA condition protocol followed the Baseline 1 probe protocol, except that only "please" was reinforced during a probe; thus, during the DRA condition all the mands listed in Table 1 were on extinction. The participant had 30 s in which to respond with "please" to obtain the item or activity. The 30-s DRA condition time period began when any of the forms listed in Table 1 occurred. That is, the behavior of the participant determined when the 30-s period began. For example, during lunch, a probe food item was placed in sight but out of reach. If the participant led the experimenter's hand toward the food (e.g., used a leading response), the experimenter would not reinforce this form, but instead would continue to make himself or herself available to fulfill a request. The participant then had 30 s in which to respond with "please." If the participant continued to respond with the forms in Table 1 for the duration of the 30 s, he or she was removed from the situation, redirected to another activity for 1 min, and then returned to the same situation. After returning, the participant could directly obtain the probe item. If the participant responded with "please" as either the initial form or as a later form within the 30-s probe, it was reinforced immediately and the probe ended. Thus, during the DRA condition as well as during all baseline probes, "please" was always reinforced and was never on extinction.

RESULTS

Ann

Figure 1 and Table 2 present the results of Ann's performance during baseline and the DRA condition. The total time, from the first baseline probe to the final DRA condition probe, excluding school vacations, was 4 months. During the first three Baseline 1 probe sets, Ann never responded with "please." Massed trial training began immediately after the Baseline 1 probes were completed. On the probe set conducted in the generalization settings during massed trial training, Ann did not respond with "please" on any probes. She required 146 trials over 14 training days to reach criterion. She maintained a 100% unprompted correct use of "please" except upon her return from a 6-week vacation. She rapidly reached criterion again, and then remained at 100% performance for the rest of the study.

After Ann reached massed trial acquisition criterion the first time (i.e., before summer school ended), the fifth baseline probe set

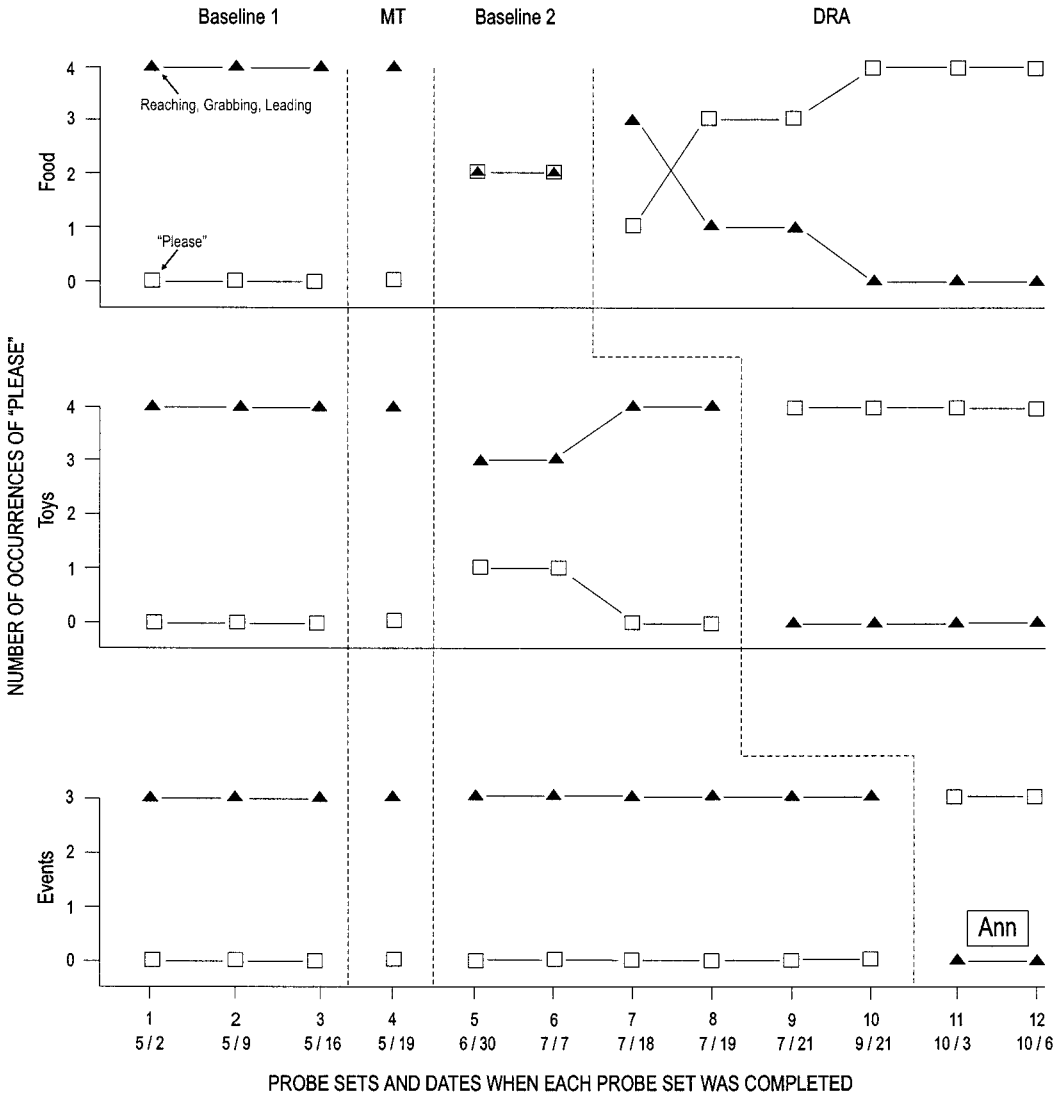

Figure 1. Number of probes in which "please" occurred for Ann. Massed trial training to acquisition criterion occurred between the first two dashed lines, and differential reinforcement of alternative behavior (DRA) occurred after the third dashed line. During Baseline 1, massed trial training (MT), and Baseline 2 (postacquisition), the first response was reinforced and the probe ended. During DRA, only "please" was reinforced.

was conducted under the same conditions as Baseline 1 probe sets (i.e., the first form Ann used during a probe was reinforced immediately and the probe ended). During this fifth probe set, Ann responded with "please" on 27% (3 of 11) of the probes. To determine whether the occurrence of her "please" response would continue to increase without further intervention, another postacquisition probe set was conducted. In this sixth probe set, Ann responded with "please" again in 27% (3 of 11) of the probes. Because the results of this probe set were consistent with those of the fifth probe set, the DRA condition was introduced in a sequential fashion across each baseline. Ann did not respond with "please" during the remaining 20 baseline condition probes.

Table 2
Percentage of Probes in Which "Please" Occurred

Participant	1	2	3	4	5	6	7	8	9	10	11	12
At any time												
Ann	0	0	0	0	27	27	9	27	64	73	100	100
Kate	0	0	0	0	0	27	45	45				
Tom	0	0	0	0	8	31	62	92	92			
As the first response												
Ann	0	0	0	0	27	27	0	18	27	36	45	45
Kate	0	0	0	0	0	9	9	0				
Tom	0	0	0	0	8	0	23	23	69			

On the seventh probe set, the DRA condition was introduced for the first time in the food category. Ann responded with "please" on only one probe (25%; 1 of 4), and that "please" was preceded by two reaches. On the remaining three DRA condition probes in the food category, she responded with an average of six old forms (range, 4 to 8) during each probe. A second DRA condition probe set was conducted in the food category because the frequency of "please" actually decreased from the postacquisition baseline phase. This second DRA condition probe set also permitted an assessment of stability in the toy category. The results of this DRA condition probe set in the food category revealed that Ann responded with "please" on 75% (3 of 4) of the probes. On 50% (2 of 4) of the probes, it was her first response.

On the ninth probe set, the DRA condition was extended to include the toy as well as the food category. Ann responded with "please" on 88% (7 of 8) of the probes. On 38% (3 of 8) of these, it was her first response. After this probe set was completed, summer school ended. When Ann returned to school and met criterion during massed trial training, a second DRA condition probe set including the food and toy categories was conducted. During this probe set, Ann responded with "please" during 100% (8 of 8) of the probes. It was her initial response on 50% (4 of 8) of the probes.

On the 11th probe set, the DRA condition was extended to include all three legs of the multiple baseline. Ann responded with "please" on 100% (11 of 11) of these probes, and it was her initial response on 45% (5 of 11) of them. On the second DRA condition probe set that included all three legs of the multiple baseline, Ann again responded with "please" on 100% (11 of 11) of the probes, and it was her initial response on 45% (5 of 11) of them. None of Ann's baseline probes (of 86) or DRA condition probes (of 46) were terminated because of problem behavior.

Kate

Figure 2 and Table 2 present the results of Kate's performance during baseline and the DRA condition. The total time, from the first baseline generalization probe to the final DRA condition probe, excluding one school vacation, was 6 months and 2 weeks. During the first three Baseline 1 probe sets, Kate did not respond with "please." Massed trial training began immediately after the Baseline 1 probes were completed. On the probe set conducted in the generalization settings during massed trial training, Kate did not respond with "please" on any probes. She required 153 trials over 20 training days to reach criterion. She maintained a 100% unprompted correct use of "please" during one massed trial maintenance session, and then summer school ended. Upon her return

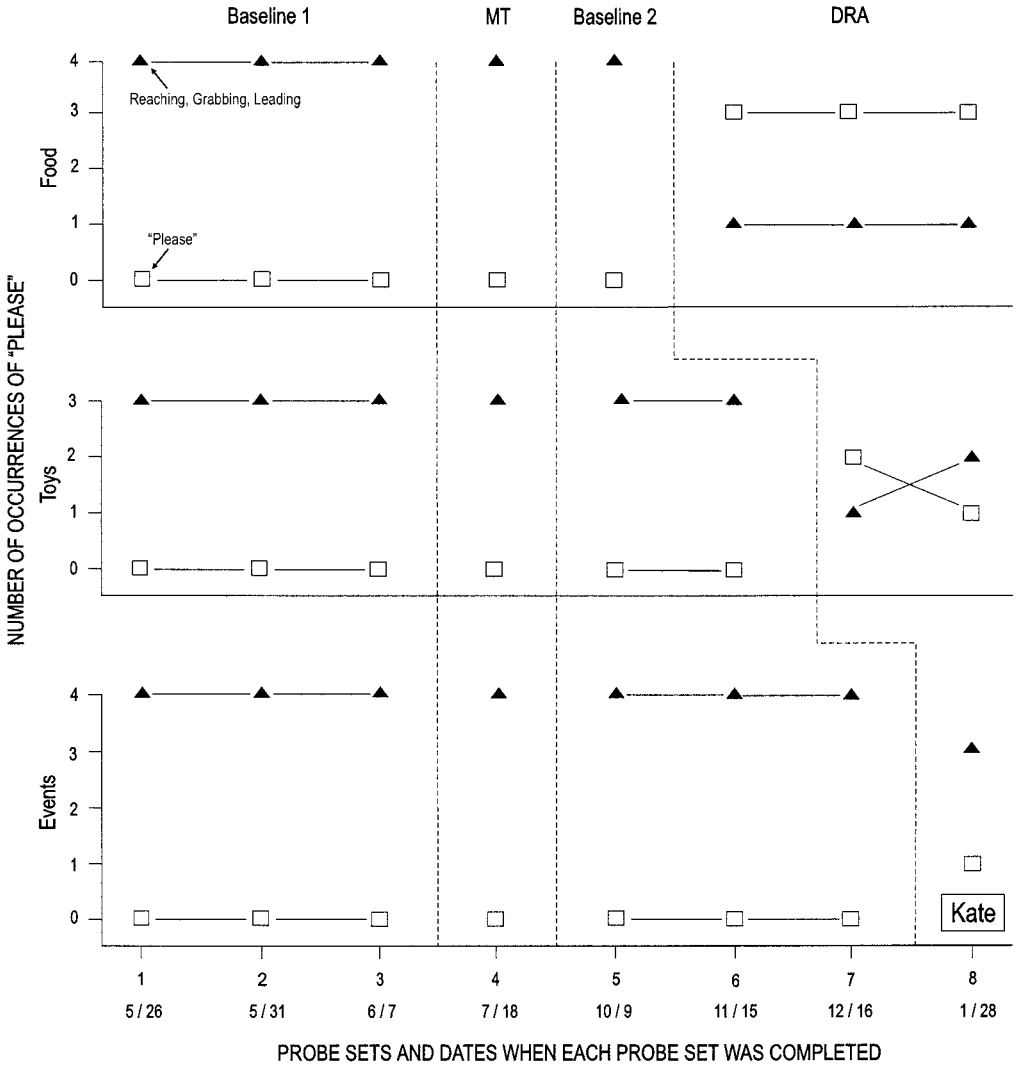

Figure 2. Number of probes in which "please" occurred for Kate. Massed trial training to acquisition criterion occurred between the first two dashed lines, and differential reinforcement of alternative behavior (DRA) occurred after the third dashed line. During Baseline 1, massed trial training (MT), and Baseline 2 (postacquisition), the first response was reinforced and the probe ended. During DRA, only "please" was reinforced.

to school 6 weeks later, Kate required 3 days of training and 28 trials to reach criterion again, and then she remained at 100% performance throughout the remaining 19 weekly maintenance training sessions.

After Kate reached criterion the second time, the fifth probe set was conducted under the same conditions as previous baseline probe sets (i.e., the first form she responded with during a probe was reinforced immediately and the probe ended). During this probe set, she did not respond with "please" on any probes. Following this fifth probe set, the DRA condition was introduced in a sequential fashion across each baseline. Kate did not respond with "please" during the remaining 11 baseline condition probes.

On the sixth probe set, the DRA condi-

tion was introduced for the first time in the food category. Kate responded with "please" on three of four probes (75%). "Please" was her first response on one probe.

On the seventh probe set, the DRA condition was extended to include the toy as well as the food category. Kate responded with "please" on 71% (5 of 7) of the probes; it was her first response on only one probe (14%). One of the probes in which "please" did not occur was terminated because of Kate's problem behavior (e.g., flopping to the floor, screaming).

On the eighth probe set, the DRA condition was extended to include all three legs of the multiple baseline. Kate responded with "please" on 45% (5 of 11) of the probes. "Please" never occurred as her initial response. Two of the probes in which "please" did not occur were terminated after 10 s because of Kate's problem behavior (e.g., knocking over furniture, screaming). A second DRA condition probe set was not conducted for two reasons. First, Kate's problem behavior was increasing during probes and, second, Kate often did not respond to probe items and events when they were available during the last probe set. Thus, the time to conduct a single probe often exceeded 2 or 3 hr. None of Kate's baseline probes (of 66) were terminated, but three (of 22; 14%) of her DRA condition probes were terminated because of problem behavior.

Tom

Figure 3 and Table 2 present the results of Tom's performance during baseline and the DRA condition. The total time, from the first baseline probe to the final DRA condition probe, was 4 months and 1 week. During the first three baseline probe sets, Tom never responded with "please." Massed trial training began immediately after the three Baseline 1 probe sets were completed. On the probe set conducted in the generalization settings during massed trial training, Tom did not respond with "please" on any probes. Tom required 424 trials over 27 training days to reach criterion. He maintained a 100% unprompted correct use of "please" throughout the remaining five weekly maintenance training sessions.

After Tom reached massed trial acquisition criterion, the fifth probe set was conducted under the same conditions as previous baseline probe sets (i.e., the first form Tom used during a probe was reinforced immediately). During this probe set, Tom responded with a simultaneous "please" and reach on one probe (of 13; 8%). Following this fifth probe set, the DRA condition was introduced in a sequential fashion across each baseline. Tom did not use "please" during the remaining 13 baseline condition probes.

On the sixth probe set, the DRA condition was introduced for the first time in the food category. Tom responded with "please" on all four probes. "Please" was never his first response. On the seventh probe set, the DRA condition was extended to include the toy as well as the food category. Tom responded with "please" on 89% (8 of 9) of the probes, and it was his first response on only three probes (33%).

On the eighth probe set, the DRA condition was extended to include all three legs of the multiple baseline. Tom responded with "please" on 92% (12 of 13) of the probes, and it was his initial response on three (23%) of them. On the second DRA condition probe set including all three legs of the multiple baseline, Tom again responded with "please" on 92% (12 of 13) of the probes. However, during this probe set, Tom's initial response was "please" on 69% (9 of 13) of the probes. None of Tom's baseline probes (of 78) or DRA condition probes (of 39) were terminated because of problem behavior.

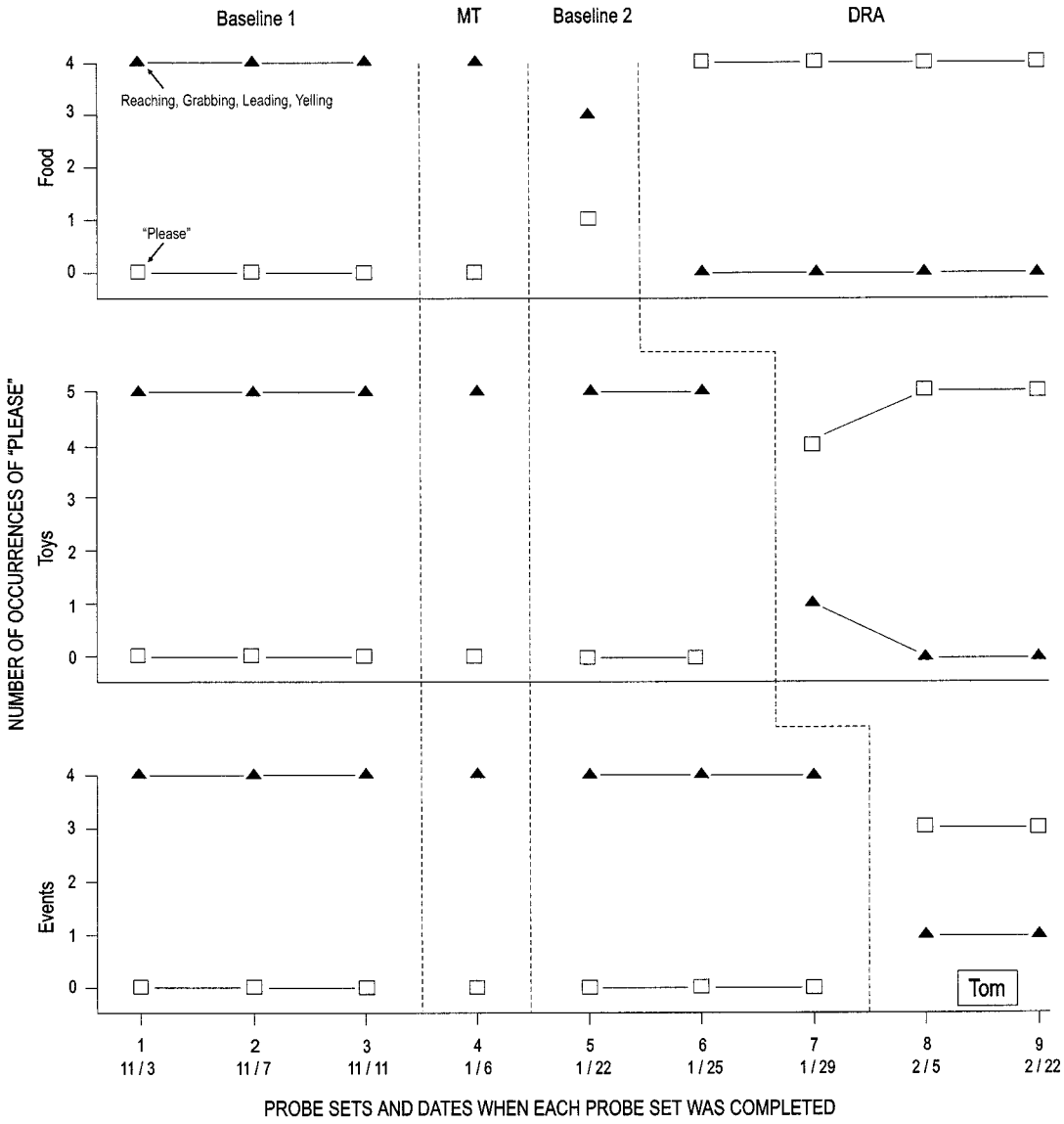

Figure 3. Number of probes in which "please" occurred for Tom. Massed trial training to acquisition criterion occurred between the first two dashed lines, and differential reinforcement of alternative behavior (DRA) occurred after the third dashed line. During Baseline 1, massed trial training (MT), and Baseline 2 (postacquisition), the first response was reinforced and the probe ended. During DRA, only "please" was reinforced.

DISCUSSION

The purpose of this study was to examine variables that may influence the generalized occurrence of a replacement mand in 3 young children with severe language delays. First, the baseline data in this study demonstrated that prior to intervention, each participant had an existing response class of mands (e.g., reaching, leading). These mands were widely generalized and evoked by a stimulus class that consisted of reinforcing but unavailable items and events. Second, when participants mastered a replacement mand (i.e., the sign for "please")

under restricted stimulus conditions, they rarely responded with it in the generalization settings. When the DRA condition (i.e., differential reinforcement of "please" plus extinction of competing mands) was introduced in the generalization settings, however, all 3 participants manifested an increase over time in the percentage of probes in which "please" occurred, with 2 of the participants demonstrating increases in "please" as their first response.

The findings of this study support those of other researchers (e.g., Carr & Durand, 1985; Lalli et al., 1995; Shirley, Iwata, Kahng, Mazaleski, & Lerman, 1997; Sprague & Horner, 1992) who showed that placing some members of a response class on extinction may increase the likelihood that alternative members of the class will be emitted. Moreover, the findings of this study revealed that a new replacement member may enter a response class even if it is acquired under restricted stimulus conditions (i.e., it is associated with only some elements of a stimulus class that control responding), but it may enter as a low-probability member of the response class. Placing existing high-probability competing members of the class on extinction, however, may produce generalization of the low-probability response. Thus, the extinction of competing forms may alter the hierarchy of response options so that initial low-probability responses become more probable (Baer, 1982; Lalli et al., 1995), even under a variety of stimulus conditions.

The multiple baseline design used in this study illustrates how existing responses continue to compete with the replacement mand. After the DRA condition was implemented, participants never responded with "please" under baseline conditions in subsequent legs, but "please" was not always the first response in the DRA condition throughout most of the study. This data pattern suggests that although the DRA condition could evoke "please" under a variety of stimulus conditions, the participants continued to respond initially with other members of their manding response class. Our analysis of this phenomenon leads to two clinical implications: First, as long as undesirable competing forms are reinforced in any setting, the new replacement behavior is less likely to occur in those settings. Thus, practitioners cannot continue reinforcing old behavior on an intermittent schedule in settings in which the new behavior is desired. Second, whenever teaching a mand, practitioners should not assume that it will be the only member of the response class. There is always a possibility that there may be some type of competing response that could interfere with the generalization of the new mand, just as all participants in this study had competing forms.

Child development research supports a response competition perspective because it has shown that the functions of requesting and rejecting (i.e., manding) emerge during the first year of life (Bates, Camaioni, & Volterra, 1975; Carpenter, Mastergeorge, & Coggins, 1983). Forms such as grabbing or reaching are early motor forms that represent these functions (Wetherby, Warren, & Reichle, 1998). Thus, when teaching a mand, it may be wise to assume that if the function exists, there are other forms that serve that function and may be widely generalized (Drasgow & Halle, 1995). Any new mand may actually be a replacement mand, and its successful generalization may depend as much on variables related to response efficiency and schedules of reinforcement (e.g., Horner & Day, 1991; Mace & Roberts, 1993) as on teaching with multiple exemplars (Stokes & Baer, 1977).

We did not analyze Kate's lower level of generalization, but several explanations seem plausible. First, Kate's massed trial training occurred in school and her generalization settings were primarily in her home, whereas

the other participants' massed trial training and probes occurred in the same setting (i.e., at school or at home). Perhaps the difference between Kate's two settings was more distinct than for the other 2 participants. Second, three of Kate's DRA condition probes were terminated because of problem behavior. Our protocol called for terminating any probes in which problem behavior occurred, and her problem behavior may have served the function of requesting (Carr, 1977; Carr & Durand, 1985). If the DRA condition probes had continued despite Kate's problem behavior, perhaps at some point she would have responded with "please." Conversely, her problem behavior may have been escape motivated because the probes became aversive to her in that they required more effortful responding (Friman & Poling, 1995). But Kate's problem behavior affected only 3 of 88 probes, and therefore it was "reinforced" in less than 4% of her probes. Third, the low or nonexistent reinforcing value of probe items may have been a major contributing factor on several DRA condition probes in which Kate did not respond with "please" (e.g., Drasgow & Halle, 1995; Egel, 1981). Evidence for this explanation includes her low average number of responses on DRA condition probes ($M = 2.8$) compared to Ann ($M = 5.2$) and Tom ($M = 8.3$), and her apparent disinterest in most probe stimuli, especially during the final DRA condition probe set.

One limitation of this study reflected by Kate's lower level of generalization was our inability to insure that probe stimuli maintained a constant and considerable reinforcing value during the duration of the study. Kate's preferences may have changed, and the original items and events that we included in her stimulus class may have lost their reinforcing value over time. A future study could address this limitation by incorporating a strategy for maintaining reinforcer value and effectiveness during the assessment of generalization (Fisher et al., 1992; Mason, McGee, Farmer-Dougan, & Risley, 1989). Another avenue for investigating the potential relationship between reinforcement value and the generalization of replacement mands is to manipulate setting events (Wahler & Fox, 1981) or establishing operations (Michael, 1982) to determine their influence on the generalization of high- and low-probability manding responses that serve either a requesting or a protesting function. We believe that this is a fruitful avenue for future exploration.

REFERENCES

Baer, D. M. (1982). The imposition of structure on behavior and the demolition of behavior structures. In H. E. Howe (Ed.), *Nebraska symposium on motivation* (pp. 217–254). Lincoln, NE: University of Nebraska Press.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1,* 91–97.

Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly, 21,* 205–226.

Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis, 1,* 175–191.

Billingsley, F. F., & Neel, R. S. (1985). Competing behaviors and their effects on skill generalization and maintenance. *Analysis and Intervention in Developmental Disabilities, 5,* 357–372.

Bricker, W., & Bricker, D. (1970). A program of language training for the severely language handicapped child. *Exceptional Children, 37,* 101–111.

Bryen, D. N., & Joyce, D. G. (1985). Language intervention with the severely handicapped: A decade of research. *Journal of Special Education, 19,* 7–39.

Carpenter, R. L., Mastergeorge, A. M., & Coggins, T. E. (1983). The acquisition of communicative intentions in infants eight to fifteen months of age. *Language and Speech, 26,* 101–116.

Carr, E. G. (1977). The motivation of self-injurious behavior: A review of some hypotheses. *Psychological Bulletin, 84,* 800–816.

Carr, E. G. (1988). Functional equivalence as a mechanism of response generalization. In R. Horner, G. Dunlap, & R. L. Koegel (Eds.), *Generalization and maintenance: Life-style changes in ap-*

plied settings (pp. 221–241). Baltimore: Paul H. Brookes.

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18,* 111–126.

Carr, E. G., & Kemp, D. C. (1989). Functional equivalence of autistic leading and communicative pointing: Analysis and treatment. *Journal of Autism and Developmental Disorders, 19,* 561–578.

Catania, A. C. (1966). Concurrent operants. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 213–270). New York: Appleton-Century-Crofts.

Catania, A. C. (1998). *Learning* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Chadsey-Rusch, J., Drasgow, E., Reinoehl, R. B., Halle, J., & Collet-Klingenberg, L. (1993). Using general-case instruction to teach spontaneous and generalized requests for assistance to learners with severe disabilities. *Journal of the Association for Persons with Severe Handicaps, 18,* 177–187.

Davison, M., & McCarthy, D. (1988). *The matching law: A research review.* Hillsdale, NJ: Erlbaum.

Drasgow, E., & Halle, J. W. (1995). Teaching social communication to young children with severe disabilities. *Topics in Early Childhood Special Education, 15,* 164–186.

Drasgow, E., Halle, J. W., Ostrosky, M. M., & Harbers, H. M. (1996). Using behavioral indication and functional communication training to establish an initial sign repertoire with a child with severe disabilities. *Topics in Early Childhood Special Education, 16,* 500–521.

Egel, A. L. (1981). Reinforcer variation: Implications for motivating developmentally disabled children. *Journal of Applied Behavior Analysis, 14,* 345–350.

Fisher, W. W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe to profound disabilities. *Journal of Applied Behavior Analysis, 25,* 491–498.

Friman, P. C., & Poling, A. (1995). Making life easier with effort: Basic and applied research on response effort. *Journal of Applied Behavior Analysis, 28,* 583–590.

Guess, D., Sailor, W., & Baer, D. M. (1974). To teach language to retarded children. In R. L. Schiefelbusch & L. L. Lloyd (Eds.), *Language perspectives: Acquisition, retardation, and intervention* (pp. 529–563). Baltimore: University Park Press.

Halle, J. W. (1982). Teaching functional language to the handicapped: An integrative model of natural environment teaching techniques. *Journal of the Association for Persons with Severe Handicaps, 4,* 29–37.

Halle, J. W., Chadsey-Rusch, J., Collet-Klingenberg, L., & Reinoehl, R. B. (1992). *Assessment manual.* Product of U.S. Department of Education Grant H086P90024.

Halle, J. W., & Holt, B. (1991). Assessing stimulus control in natural settings: An analysis of stimuli that acquire control during training. *Journal of Applied Behavior Analysis, 24,* 579–589.

Halle, J. W., Marshall, A. M., & Spradlin, J. E. (1979). Time delay: A technique to increase language use and facilitate generalization in retarded children. *Journal of Applied Behavior Analysis, 12,* 431–439.

Harris, S. (1975). Teaching language to nonverbal children: With an emphasis on problems of generalization. *Psychological Bulletin, 82,* 565–580.

Hart, B., & Rogers-Warren, A. K. (1978). A milieu approach to teaching language. In R. L. Schiefelbusch (Ed.), *Language intervention strategies* (pp. 193–236). Baltimore: University Park Press.

Hedrick, D. L., Prather, E. M., & Tobin, A. R. (1984). *Sequenced inventory of communication development* (rev. ed.). Seattle: University of Washington Press.

Hemmeter, M. L., Ault, M. J., Collins, B. C., & Meyer, S. (1996). The effects of teacher-implemented language instruction within free time activities. *Education and Training in Mental Retardation and Developmental Disabilities, 31,* 203–212.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13,* 243–266.

Horner, R. H., Bellamy, G. T., & Colvin, G. T. (1984). Responding in the presence of nontrained stimuli: Implications of generalization error patterns. *Journal of the Association for Persons with Severe Handicaps, 9,* 287–295.

Horner, R. H., & Day, H. M. (1991). The effects of response efficiency on functionally equivalent competing behaviors. *Journal of Applied Behavior Analysis, 24,* 719–732.

Johnston, J. M., & Pennypacker, H. S. (1993). *Strategies and tactics of behavioral research* (2nd ed.). Hillsdale, NJ: Erlbaum.

Kaczmarek, L. A. (1990). Teaching spontaneous language to individuals with severe handicaps: A matrix model. *Journal of the Association for Persons with Severe Handicaps, 15,* 160–169.

Kaiser, A. P., Yoder, P. J., & Keetz, A. (1992). Evaluating milieu teaching. In S. F. Warren & J. Reichle (Eds.), *Causes and effects in communication and language intervention* (pp. 9–47). Baltimore: Paul H. Brookes.

Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings.* New York: Oxford University Press.

Koegel, R. L., O'Dell, M. C., & Koegel, L. K. (1987). A natural language teaching paradigm for nonverbal autistic children. *Journal of Autism and Developmental Disorders, 17,* 187–200.

Lalli, J. S., Mace, F. C., Wohn, T., & Livezey, K.

(1995). Identification and modification of a response-class hierarchy. *Journal of Applied Behavior Analysis, 28,* 551–559.
Losardo, A., & Bricker, D. (1994). Activity-based intervention and direct instruction: A comparison study. *American Journal on Mental Retardation, 98,* 744–765.
Mace, F. C., & Roberts, M. L. (1993). Factors affecting selection of behavioral interventions. In J. Reichle & D. P. Wacker (Eds.), *Communicative alternatives to challenging behavior: Integrating functional assessment and intervention strategies* (pp. 113–133). Baltimore: Paul H. Brookes.
Mason, S. A., McGee, G. G., Farmer-Dougan, V., & Risley, T. R. (1989). A practical strategy for ongoing reinforcer assessment. *Journal of Applied Behavior Analysis, 22,* 171–179.
Michael, J. (1982). Distinguishing between the discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior, 37,* 149–155.
Parrish, J. M., Cataldo, M. F., Kolko, D. J., Neef, N. A., & Egel, A. L. (1986). Experimental analysis of response covariation among compliant and inappropriate behaviors. *Journal of Applied Behavior Analysis, 19,* 241–254.
Reichle, J., & Sigafoos, J. (1991). Establishing spontaneity and generalization. In J. Reichle, J. York, & J. Sigafoos (Eds.), *Implementing augmentative and alternative communication: Strategies for learners with severe disabilities* (pp. 157–171). Baltimore: Paul H. Brookes.
Schiefelbusch, R. (1978). *Language intervention strategies.* Baltimore: University Park Press.
Schiefelbusch, R., & Lloyd, L. (1974). *Language perspectives: Acquisition, retardation, and intervention.* Baltimore: University Park Press.
Shirley, M. J., Iwata, B. A., Kahng, S. W., Mazaleski, J. L., & Lerman, D. C. (1997). Does functional communication training compete with ongoing contingencies of reinforcement? An analysis during response acquisition and maintenance. *Journal of Applied Behavior Analysis, 30,* 93–104.
Sprague, J. R., & Horner, R. H. (1992). Covariation within functional response classes: Implications for treatment of severe problem behavior. *Journal of Applied Behavior Analysis, 25,* 735–745.
Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis, 10,* 349–367.
Wahler, R. G., & Fox, J. J. (1981). Setting events in applied behavior analysis: Toward a conceptual and methodological expansion. *Journal of Applied Behavior Analysis, 14,* 327–338.
Warren, S. F., & Rogers-Warren, A. K. (1985). *Teaching functional language: Generalization and maintenance of language skills.* Austin, TX: Pro-Ed.
Wetherby, A. M., Warren, S. F., & Reichle, J. (Eds.). (1998). *Transitions in prelinguistic communication.* Baltimore: Paul H. Brookes.

Received September 2, 1997
Initial editorial decision October 31, 1997
Final acceptance April 8, 1998
Action Editor, Robert H. Horner

STUDY QUESTIONS

1. What reason do the authors provide as one explanation for limited generalization of newly acquired communication responses?

2. What were the typical mands exhibited by the participants, how were these mands identified, and what replacement response was chosen for training?

3. Describe the essential components of the baseline protocol.

4. Describe the massed trial training protocol. How was the possibility of response chaining addressed? Do any results suggest that chaining may have occurred?

5. What was the purpose of the probes during and immediately following the massed trial training phase (Baseline 2)? What results were obtained during these probes?

6. Describe the contingencies in effect for previous mand forms, newly taught mands, and inappropriate behavior during the DRA condition.

7. Summarize the results obtained during the DRA condition in terms of the usage of (a) "please" and (b) previous mand forms as first and as subsequent responses. To what extent do these data reflect generalization?

8. It was suggested that Kate's somewhat limited use of "please" during the DRA condition may have been because the preferred items lost their reinforcing value (i.e., satiation occurred). What feature of her data (see Figure 2) might suggest that the items maintained their reinforcing value?

Questions prepared by Gregory Hanley and Jana Lindberg, The University of Florida